International Journal of Education and Language Studies

ISSN: 2791-9323

Khalid Ahmed Hassan AHMED [1]

THEORY AND PRACTICE IN ENGLISH LANGUAGE TESTING

Istanbul / Türkiye

p. 13-22

Article Information

Article Type: Research Article

This article was checked by **iThenticate** No **plagiarism** detected

Article History

Received: 02/05/2021
Accepted: 20/05/2021
published: 01/06/2021

Abstract:

Success in language teaching and learning is a multidisciplinary process in which language testing is an important pillar of success beside the other surrounding factors which range from the teaching learning context to the learners and teachers' personal and educational factors. The aim of the present study is to identify to what extent there is a harmony between the teaching learning context on one hand and the theories and techniques of language testing on the other, and whether there is a matching between these theories of language testing and the practicality of these theories. We believe that there is a great mismatch between theories and practice of language testing as a result of the practices and the necessary expertise and knowledge needed in testing.

This situation affects, to a great extent, the experience of language teaching learning process. This paper is a qualitative historical survey on language teaching learning context and theories of testing and language testing in particular. To process the study some related studies and literature will be explored, and some instances of language testing will be used to support this theoretical framework and to confirm our hypothesis that there is a gap between what is taught and what is examined. The study will include its importance, the methodological procedures that will be followed to attain the findings, recommendations and suggestions for future studies to let discussion in this area furthered in the future.

Key words: Theory, Practice, Language Testing, Validity, Reliability, Mismatch, Testing Approaches, Testing Techniques, Learners' Orientation, Motivation, Anxiety, Learning Materials.

http://dx.doi.org/10.47832/2791-9323.2-2.2

[1] Dr. , Omdurman Islamic University, Sudan, khalidsoon@hotmail.com, https://orcid.org/0000-0003-1115-1114

Introduction

Teaching a language is an integrated environment where variety of factors play essential roles in making that experience a successful voyage of teaching and learning. Language testing is the measurement through which success or failure of this teaching and learning experience can be assessed. Testing should be constructed and set up through the best theories and techniques that are normally designed to guarantee such a good mastery of a test design. However, teaching and learning languages experiences, such as the other scientific practices, may also undergo some kind of a mismatch between theory and practice in the whole experience, or in particular a mismatch between theory and practice in the set up and design of testing.

Variety of factors may integrate to let such a mismatch between theory and practice in language testing take place. In a part the factors may be related to the approaches, methods and techniques of language testing. That is to say practitioners of language teaching and learning may not have good training or knowledge of language testing methodology; therefore, they may set up their tests in a way that does not confine with the beast ways of designing good language tests that actually measure what should be measured in terms of testing validity, or in terms of getting consistent learners' scoring in terms of test reliability. Testing in general is believed to be a motivating factor of learning when it works as a director or guide for the teaching and learning process. On the other side the mismatch between theory and practice in testing may result of the language teaching learning process mainly in cases where there is no concern about providing better training facilities to upgrade the teaching learning experience. That is to say untrained and unqualified language teachers are expected to have more mismatch between theory and practice in setting up their language tests.

To discuss the matter of mismatch between theory and practice in English language testing in particular, we need to discuss in some details all the theoretical matters concerning the approaches and theories of testing and to correlate that with English language teaching and learning methodology. Therefore, the theoretical framework and the review of the related literature below will give us a good stand for such a discussion.

The Importance of the Study

This study is Important as it focuses on an important aspect of language teaching and learning methodology, that is the relationship between theory and practice in English language testing and in particular the mismatch between theory and practice in English language testing as an expected outcome of this relationship between testing as a theory and its real practices. The study is important as it tries to provide practical suggestions for avoiding any inconveniences or gaps between teaching and testing because the two should be in a harmony for a successful teaching learning experience. It is also important as a result of the integration between the previous research projects and this present study. The study is also important as a result of the accompanying of a personal experience of someone who spends years in teaching and testing, and it is known that experience is a fruitful angle from which theories and their practices may be put in the right order.

Methodology

The study is qualitative approach, content analysis and a historical survey over the related literature in its processing, discussion, findings and conclusion. For that end a review, discussion and correlation among the previous related studies will be undergone together with our own observations and experience as a practitioner in the field of English language teaching and testing. In the first part of the discussion the theoretical frame work of the study will be outlined to include language testing approaches, methodology and techniques. The theoretical frame work of testing will be correlated with the methodology of English language teaching and learning in order to detect the sources of the mismatch between theory and practice in English language testing. The theoretical framework will further be supported by the review of the related literature review concerning aspects of language testing in particular and the expected sources of mismatch between theory and practice.

Procedures

As it has been stated within the methodology of this study that this study will be a correlation between the theoretical frame of testing, the previous related literature and our own observations of teaching and testing practice, within the theoretical frame work and the related studies these concepts will be intensively discussed followed with comments and discussions to let the study be processed. Some examples of tests will be shown to reflect how these examples confirm the central hypothesis that there is a mismatch between theory and practice in English language testing. The theoretical framework and the related literature will be followed with the discussion and the correlation between the theoretical frame work and the related studies in order to come out with the main findings, suggestions and recommendations for further future studies out of this present study.

The theoretical frame work and some related studies

As it has been stated within the methodology and the procedures for this study the analysis, discussion and interpretation of the processing of the study will be based on the general theoretical frame work of testing and the previous related literature which tackle aspects related to this present study, this section will be devoted to these theoretical aspects of the study. First we need to set up the theoretical aspects of testing regarding the classifications of the types of tests and the testing techniques, and how tests are measured and evaluated in order to prepare the grounds for how the mismatch between theory and practice happens as a result of the misunderstanding of testing conceptualization and administration.

As a starting point Johnson and Johnson (1999: 198) make a correlation between language teaching and language testing by saying that language testing is made up of a variety of techniques and regulations which are not found in language teaching, but the two interact somewhere for the general assessment of learners' language performance and achievement. The same idea has also been maintained by Ellis et al (2009: 337) when they cited Lantolf's (2008) argument about the correlation between the implicit and explicit knowledge between language testing and teaching that lead to educational benefits in studying a language. Davies and Elder (2004: 414) and (Hawkey (2000: 8) express this relationship between teaching and testing by the term 'backwash' which indicates the influence of teaching on testing or testing on teaching, and it will be either beneficial or positive when the relationship leads to success and a negative backwash if it is not a good successful teaching

testing experience. In a more general theoretical aspect it is a well known practice that tests are subdivided into two major types according to the setting up of the tests and the scoring processes. These two general types of tests are subjective and objective tests. The subjective test is the type of testing that is greatly affected by the scorer's personal views, while the objective test is the type of test that is governed by the type of the test and its techniques of designing and setting up. However, we expect that the mismatch between theory and practice is mainly expected to be within the objective types of tests which are generally identified as more complicated in terms of their setting up and design, and which are simple and easy in terms of their marking and correction, but the subjective tests are simple and easy in their setting up while they are more complicated in terms of their marking and correction.

Hughes (2003: 11 - 23) explains the different types of tests and their theoretical frame works in the following summary of each type and its theoretical framework and theoretical requirements: first he outlines major types of language testing such as proficiency tests which evaluate candidates overall ability in a language without linking that to a particular course of study. Achievement tests which are directly connected with particular programmes or courses of study, and with which language teachers are directly concerned with in terms of their design or administration. Diagnostic tests which are designed for detecting students' strengths and weaknesses in order to plan for further teaching and learning requirements or remedial processes. Placement tests which are the types of tests that help in subdividing learners within their liable group level. Direct and indirect tests are the tests in which we try to measure students abilities within either directly related skills such as the writing of composition in the case of the direct tests, and in the cases of indirect testing we run after assessing infinite number of skills or sets of skills. Discrete point versus integrative testing which imply the testing of just a single item of learning at a time in the first type, or the combination of variety of items in the case of integrative testing. Norm – referenced versus criterion – referenced testing which simply indicate that in the norm – referenced case we rate a certain student's performance with the performance of his/her peers; while in the second case we will be interested in the overall performance of a learner and what he/she can do with the target language. Then Hughes (ibid) gives the same idea mentioned above about the subjective and objective language testing. Then he mentioned computer adaptive testing which is the type of testing that is designed to address the students' abilities by first giving all the students tests with average standardized items, the students who get high scores will be given more difficult itemed tests, and those who show low standards will be given easier items. Then Hughes (ibid) ends up with communicative language testing which implies the testing of the students' abilities to use the target language communicatively in all the language skills.

Hughes (2003: 75) mentions the techniques by which testing can be designed. These techniques include: Multiple choice items which include a stem as the central structural unit such as "Enid has been here ___________________ half an hour." The stem is followed with some options; one of which is the correct choice and the others are distracters such as : A. during B. for C. while D. since. After that Hughes (ibid) mentions the other techniques such as yes/ no questions and true and false, short answer items, and gap filling items. The central points within these techniques that they should be highly designed in terms of their language construction and the overall setting of a test to ensure test validity in terms of testing what should be tested, and test reliability in terms of students' consistency of scores in case of administering the tests more than once. However, we believe that

the design and the language structure of the test will be one of the factors that affect the relationship between testing and testing and language theories, the thing that may result in what is called a negative backwash between teaching and testing as will be discussed below.

In tracing the history of language testing Madsen (1983: 5 – 6) said that language testing in the nineteenth and the beginnings of the twentieth century used to be commonly influenced by teachers' intuitive personal impressions in the form of subjective tests that test formal issues of language rather than language use. Then he classified the intervention of objectivity in testing by a scientific trend that comes after the intuitive subjective era. These subjective tests rely on students' labeling parts of sentences and the memorization of grammatical rules or lexical items, essay writing, précis, dictations, open ended answers, etc. He also classified the teachers who used to design such types of tests as un trained or unqualified teachers. However, after the objective testing era the trend was described as a scientific movement in which specialized language experts should measure language use in an era of communicative language use.

For Alderson et al (1995: 128 – 9) a thorough training is needed mainly for the marking of subjective tests because training will give the examiners the confidence needed to encounter possibility of unseen problems when designing a test. The expected problems that may affect the examiners rating and reliability may either be related to the rating scheme of marking or may be social or personal with regard to the examiner himself. The same training is also needed in cases of objective tests, but the demand is not as great as in the case of the subjective test, and the marker's reliability may just be of checking whether he/she follows the designed marking scheme and the rating of the scores. The marker's reliability is measured by either he/she shows more or less the same scores in different administrations of the same students, and this is called 'intra – rater – reliability' and if not it will be classified as 'inter – rater – reliability'.

According to Bachman and Cohen (2001: 91) the test items designed to test reading comprehension may reflect false interpretation as a result of what they classified as 'test – wise techniques' by which the candidates find ways to get the correct answer without even understanding the reading text. This phenomenon may be a common practice among teachers or students when they try to get ready for tests or examinations; that is to say teachers may try to train students to develop such 'test – wise' techniques without greater interest in teaching the language for a better competence and performance. We think that this practice of developing techniques to pass examinations presents a major component of the mismatch between theory and practice in testing. However, Hutauruk (2015: 25) in explaining the concept of communicative language testing gives a picture of how matching between theory and practice can be attained in such types of tests because the real communicative test should reflect the real learner's knowledge of the forms of the language and how to use this knowledge of the forms of the language in real life contexts through his/her performance of language items. In still another dimension of the relation between theory and practice in language testing Davies (2007: 29) relates language testing research as an activity among applied linguists by saying that such a type of research runs after increasing our knowledge of language learning and gives us glimpses about even new domains of still new language uses. Davies (2007: 84) goes on discussing aspects of language testing, and proficiency tests in particular, by saying that testing is a genuine part of the activity of language teaching because the prime goals of a language teacher are skills and knowledge. Testing is required to attain a native language speaker performance.

In an article about linking teaching with testing Hancock (2006: 12) concludes his discussion by saying that the current concern of foreign or second language instructors should be directed to assessing or testing their students on how they can use the language instead of just assessing the mere

knowledge of that target language. In other words teaching and testing should lead to learners' actual use of the target language. According to Sárosdy et al (2006: 136) testing and teaching should be perceived as partners, that is to say testing should activate good teaching that is reflected in good testing, and above all testing can be considered as a remedy if teaching happens to be of bad quality. Testing is needed by learners to be sure about their language achievement. Then Sárosdy et al (2006: 137) outline the main types of tests in the following list:

" - measure overall language proficiency independent of any language course that the candidates may have attended - proficiency test

- measure the students' achievement on a completed course of studies - achievement test

- measure how much the students have learnt of the recently taught material progress test

- diagnose students' strengths and weaknesses in language knowledge and use, to find out what they know and what they do not know

-diagnostic test - assist placement of students by identifying the stage or part of a teaching programme which is the most appropriate to the level of their proficiency – placement test

- identify general abilities, to find out who is to be good at learning languages – aptitude test"

Fulcher and Davidson (2012: 24 -5) discuss the issue of theory and practice in language testing under a subtitle that summarizes this relationship as usefulness of language testing by mentioning Bachman and Palmer's ideas (1996) about test validity and that test validity should be evaluated in terms of usefulness with regard to whether the test responded to the set up uses for which it has been designed. Then they outlined six measurements to be taken in the evaluation of the test as useful or not. These six criteria are:

" Reliability: The consistency of test performance that is reflected in the test scores.

. Construct validity: The appropriateness of the inferences about test takers' ability made on the basis of test scores.

Authenticity: The correspondence between the test tasks and the tasks students will engage in outside the test context.

Interactiveness: The extent of the involvement of test takers' engagement in completing the test tasks using the same capacities they would use in other contexts.

Impact: Positive consequences the test might have on individuals and institutions within the immediate test setting and beyond.

Practicality: The extent to which resources support the design and development of the test as well as its ongoing use."

These six points give the gist of the matching between theory and practice of language testing. That is to say the test responds to its usefulness in terms of criteria related to its theoretical aspects of its setting up and design, and at the same time to help in attaining the required educational and social requirements of the test takers in the community.

Discussion, Results, Conclusion and Recommendations

As it has been stated within the methodology and the procedures of this present study the discussion and the analysis will mainly be based on the reviewed theoretical frame work and the specialized literature of testing in addition to the personal observations regarding testing design and administration. It is clear that we have strong evidence to support that there is a gap between testing as a theory and the actual testing practices that we observe. The reasons behind this gap may either be personal with regard to the practitioners of testing themselves, or purely theoretical with regard to the conceptualization of testing and its processing, the points which will be analyzed and discussed below.

With regard to the personal factors that affect the relationship between theory and practice, mainly if we take the Sudanese community as an example, we can say that variety of factors can be of great influences in designing and administering testing. In certain cases it is found that the appointed teacher who is going to be responsible of teaching and testing may lack the essential requirements of training to be a teacher and hence a language tester. Therefore, what he does to practice teaching and testing may just be his/her experience as a former student and how he/she used to be taught or tested away of whether that was appropriate to be followed or not. Moreover, teachers processes of intake to be appointed as teachers may just be getting any kind of certification without consideration to the major specialization, graduate of colleges of education. Moreover, the profession of teaching itself has become unpromising and discouraging profession in terms of the financial or even contextual factors. Within the theoretical frame work we discussed the intervention of the teacher's problems in the course of his/her teaching experience that may lead to a negative backwash in the relationship between teaching and testing. Therefore, a healthy educational context is needed to grant the least necessities of a good positive backwash that makes a good harmony between teaching and testing in creating and graduating competent and fluent language students.

With regard to the other side, the mere theoretical factors, of the problem of the mismatch between theory and practice in language testing variety of factors may intervene to complicate matters. Although the personal factors and the theoretical factors of the mismatch between theory and practice in language testing may not be separable, it is better to be discussed a lone without affecting them with the personal factors because they may be problems to be detected everywhere due to the real comprehension and the real understanding of language testing. These problems are mainly connected with the good comprehension of language testing as a theory. That is to say if testing is practiced without a real understanding of its conceptualization it will lead to a negative backwash and there will be a lack of harmony between what is taught and what is tested. Therefore, we have reasons to say that test validity and reliability will not be attained due to this gap between theory and practice. The first practices to be observed will mainly be those of testing techniques as it has been explained above within the theoretical framework of this study. For example in multiple choice questions there should be a consistency between the stem, the correct option and the distracters; theoretically speaking the distracters should be correct options within the context of the stem but the correct option is the only correct choice as can be seen within the example mentioned above within the theoretical framework to explain the multiple choice questions. In that question all the distracters can be possible answers, but 'for' is the best option and answer according to the grammatical context of the present perfect tense. In one of the Sudanese Certificate Examinations the following stem of a multiple choice question was given to the students in the nineties of the previous century " Mandela ____________ a Nobel prize since 1986." the correct option for this stem was 'has been given" according to the context of the present perfect. However, this multiple choice item is one of the problems that we encounter in designing such types of questions; that is to say such a type of stems is faulty because the correct stem should be "Mandela _________________ a

Nobel Prize in 1986." and the correct option should be "was given" which is a context of past simple in the passive voice. Such examples of badly structured questions will definitely affect the validity of the test. The problems of tests designing will also be of great influences on tests validity the thing that will affect the balanced distribution of the teaching learning items of the syllabus, and this unbalanced distribution will definitely affect the processes of teaching and learning the target language. Therefore, we may expect an extended list of taught language items to be given a very limited portion of testing items and other very limited language items to be given a lot of emphasis in the testing items. In other words the design of the test is not built upon a proportion between the percentage given to the teaching and that is given to testing. In order to design tests according to the required criteria of testing first, there should be an ongoing training in language testing to keep away from any drawbacks of unwise practices of testing theories. The training will provide the language teachers and testers with the confidence needed to face any kind of problems whether the foreseen or the unforeseen ones. Training will also provide the practitioners with the knowledge needed to be in touch with the renewed issues in language testing.

In addition to the wrong designing of tests, there is also the problem of solving students' weaknesses in English language. The readily existing solution for such weaknesses will be testing. That is to say teachers try to just help students through training them in how to pass examinations instead of teaching them the language that will help them to be competent and proficient English language users. The students themselves adapt their strategies of learning English language to that examination oriented type of teaching and learning.

Tests should be designed and administered to attain the following teaching and learning requirements for a positive backwash between teaching and testing. These requirements, as it has been stated within the theoretical framework, are: first students' test ' Reliability' that is reflected in the test scores. Second, ' Construct validity' which implies the interpretations of the test scores as representatives of the test scorers attained language abilities. Third, 'Authenticity' which implies the matching between what is learned and what is expected to be used and to be encountered outside the domain of the language teaching, learning testing context. Fourth, the 'Interactiveness' which is the use of more or less the same capabilities of completing the test tasks in other similar contexts. Fifth, 'Impact' which implies the positive consequences that the test will have on both the test takers and their institutions within the test setting or beyond the test setting. And finally 'Practicality' which means the support that will be directed to the construction of the test and its development and continuity. All these points indicate that the process of test settings, design, support, administration, and interpretation should be integrated processes to reflect some kind of consistency between the theoretical aspects and the practical aspects of testing for a fruitful and a positive language teaching and learning experience.

Summary of the Main Findings and Results

In the light of the discussion and the interpretations mentioned above to process this study and to link it with its methodology and procedures, we feel that there are strong evidences to support the following obtained findings and results:

1. Testing practices reflect a great mismatch between theories and practices in the designing and administration of test.
2. Personal social teachers' factors have negative consequences on testing designing.
3. Lack of enough experience and testing knowledge lead to badly designed and structured tests.
4. Students' weaknesses lead to concentration on testing more than on teaching.
5. Students adapt themselves to the testing oriented mode of teaching and learning.

Recommendations and Suggestions for Further Studies

According to the nature of this study that has been designed to detect the influence of theory and practice in language testing and its practices on supporting teaching and learning, we feel that this present study should be replicated in other future studies in order to carry on further discussions on aspects that are not given greater concern in this study. Therefore, the following aspects will help this topic to be taken in other future studies, these include:

1. Incorporation of testing theory and practice in syllabus design.
2. The impact of testers teaching learning experience on testing.
3. Developing approaches to neutralize testers personal and social factors on testing.
4. Adopting teaching and learning modes to replace testing modes practices.

References

Alderson, J. Charles, Caroline Clapham and Dianne Wall (1995) Language Test Construction and Evaluation. Cambridge University Press.

Bachman, Lyle F. and Cohen, Andrew D. (2001) Cambridge University Press.

Davies, Alan and Elder, Catherine (2004) The Handbook of Applied Linguistics. Blackwell Publishing.

Davies, Alan (2007) An Introduction to Applied Linguistics From Practice to Theory. Edinburgh University Press

Ellis, Rod, Shawn Loewen, Catherine Elder, Rosemary Erlam, Jenefer Philp and Hayo Reinders (2009) Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching. MULTILINGUAL MATTERS Bristol • Buffalo • Toronto.

Fulcher , Glenn and Davidson , Fred (2012) The Routledge Handbook of Language Testing. Routledge Taylor & Francis Group London and New York.

Hancock, Charles R. (2006) Language Teaching and Language Testing: a way and ways to make appropriate connections. The Ohio State University (Columbus - Ohio - USA) / hancock.2@osu.edu

Hawkey, Roger (2000) Impact Theory and Practice. Cambridge University Press.

Hughes, Arthur (2003) Testing for Language Teachers. Cambridge University Press.

Hutauruk, Bertaria Sohnata (2015) A Compilation Material of English Language Testing. UNIVERSITAS HKBP NOMMENSEN PEMATANGSIANTAR.

Johnson, Keith and Johnson, Helen (1999) Encyclopedic Dictionary of Applied Linguistics A Handbook for Language Teaching. Blackwell.

Madsen, Harold S. (1983) Techniques in Testing. Oxford University Press.

Sárosdy, Judit, Tamás Farczádi Bencze, Zoltán Poór, and Marianna Vadnay (2006) Applied Linguistics I for BA Students in English. Bölcsész Konzorcium.